Číselné algoritmy

Index Calculus; Faktorizace pomocí GDLP

31. 5. 2021

Faktorizace pomocí zobecněného DLP

- Na vstupu je složené N, chceme najít jeho netriviální dělitel.
- K dispozici je orákulum, které řeší problém diskrétního logaritmu v cyklických podgrupách $\mathbb{Z}_N^*$.
- Vzhledem k povaze problému nepředpokládáme znalost řádu grupy.
- Postup podle článku E. Bach: *Discrete logarithm and factoring* (1984)

Zobecněný problém diskrétního logaritmu

Máme konečnou grupu G a prvky $g, h \in G$ takové, že $h \in \langle g \rangle$.
Chceme určit $n \in \mathbb{N}$ takové, že $g^n = h$.
Přestože nepředpokládáme znalost $|G|$ ani $o(g)$, můžeme přesto požadovat omezení na velikost n například v závislosti na odhadu $|G|$.
V našem příkladě je $G = \mathbb{Z}_N^*$, bez znalosti faktorizace N nepředpokládáme znalost $|G|$, ale víme, že $|G| \leq N$.
Pro zobecněný problém diskrétního logaritmu pak dává smysl přidat podmínku $n \leq N^e$, kde e je zvolená konstanta.

Značení

Předpokládejme, že $N = p_1^{e_1} p_2^{e_2} \cdots p_k^{e_k}$,
$2 < p_1 < p_2 < \cdots < p_k \in \mathbb{P}$ a $e_1, e_2, \ldots, e_k \in \mathbb{N}$.
Tedy $\mathbb{Z}_N^* \simeq \mathbb{Z}_{p_1^{e_1}}^* \times \mathbb{Z}_{p_2^{e_2}}^* \times \cdots \times \mathbb{Z}_{p_k^{e_k}}^*$, $|\mathbb{Z}_{p_i^{e_i}}^*| = p_i^{e_i - 1}(p_i - 1)$.
$|\mathbb{Z}_N^*| = \varphi(N) = p_1^{e_1 - 1}(p_1 - 1) \cdots p_k^{e_k - 1}(p_k - 1)$.
Vidíme, že grupa $\mathbb{Z}_N^*$ obsahuje $2^k - 1$ prvků řádu 2, pro $k \geq 2$ tedy jistě nejde o cyklickou grupu.
Označme $\lambda = \mathrm{NSN}(|\mathbb{Z}_{p_1^{e_1}}^*|, |\mathbb{Z}_{p_2^{e_2}}^*|, \ldots, |\mathbb{Z}_{p_k^{e_k}}^*|)$.

Tvrzení

Pro každé $g \in \mathbb{Z}_N^*$ je $g^\lambda \equiv 1 \pmod{N}$.

Dále označme

$$K := \{a \in \mathbb{Z}_N^* \mid a^{\lambda/2} \equiv \pm 1 \pmod{N}\}$$

Vlastnosti množiny K

Snadno ověříme, že K je podgrupa $\mathbb{Z}_N^*$.

Tvrzení

Pokud $k \geq 2$ (tedy N je dělitelné dvěma různými prvočísly), pak $K \neq \mathbb{Z}_N^*$.

Důkaz: Pro $a \in \mathbb{Z}_N$ máme $a \in K$ právě když $a^{\lambda/2} \equiv \pm 1 \pmod{N}$.
To je ekvivaletní jednomu z následujících výroků:

1. Pro každé $1 \leq i \leq k$ je $a^{\lambda/2} \equiv 1 \pmod{p_i^{e_i}}$
2. Pro každé $1 \leq i \leq k$ je $a^{\lambda/2} \equiv -1 \pmod{p_i^{e_i}}$

Vzhledem k definici λ je $a^{\lambda} \equiv 1 \pmod{p_i^{e_i}}$ pro každé $1 \leq i \leq k$.
Protože je každá z grup $\mathbb{Z}_{p_i^{e_i}}^*$ cyklická, obsahuje právě jeden prvek řádu 2 (konkrétně $p_i^{e_i} - 1$).
Takže pro každé $1 \leq i \leq k$ je $a^{|\mathbb{Z}_{p_i^{e_i}}^*|/2} \equiv \pm 1 \pmod{p_i^{e_i}}$.

Dokončení důkazu:

Ukážeme, že také platí $a^{\lambda/2} \equiv \pm 1 \pmod{p_i^{e_i}}$ pro každé $1 \leq i \leq k$.
Pokud je $|\mathbb{Z}^*_{p_i^{e_i}}|$ dělitel $\lambda/2$, musí být $a^{\lambda/2} \equiv 1 \pmod{p_i^{e_i}}$.
Naopak, pokud je $v_2(\lambda) = v_2(|\mathbb{Z}^*_{p_i{}^{e_i}}|)$, je $\lambda = |\mathbb{Z}^*_{p_i^{e_i}}|\ell$, kde ℓ je liché číslo. Pokud bude prvek $a \in \mathbb{Z}^*_N$ splňovat $a^{|\mathbb{Z}^*_{p_i^{e_i}}|/2} \equiv -1 \pmod{p_i^{e_i}}$ bude též $a^{\lambda/2} \equiv -1 \pmod{p_i^{e_i}}$.
Předpokládejme, že $v_2(|\mathbb{Z}^*_{p_1^{e_1}}|) \geq v_2(|\mathbb{Z}^*_{p_i^{e_i}}|)$ pro $i \in \{1, \ldots, k\}$.
Podle CRT existuje právě jedno $a \in \mathbb{Z}^*_N$, pro které je $a \bmod p_1^{e_1}$ generátor grupy $\mathbb{Z}^*_{p_1^{e_1}}$ a $a \bmod p_i^{e_i} = 1$ pro $i \geq 2$.
Pak $a \in \mathbb{Z}^*_N \setminus K$.

$K \neq \mathbb{Z}_N^*$

Důsledek

Náhodně vybraný $a \in \mathbb{Z}_N^*$ *neleží v* K *s pravděpodobností alespoň* $\frac{1}{2}$

Důkaz: $\frac{|K|}{|\mathbb{Z}_N^*|} = \frac{1}{[\mathbb{Z}_N^* : K]} \leq \frac{1}{2}$. Proto $1 - \frac{|K|}{|\mathbb{Z}_N^*|} \geq \frac{1}{2}$.

Faktorizace pomocí prvků z $\mathbb{Z}_N^* \setminus K$

Necht' $a \in \mathbb{Z}_N^* \setminus K$. Prohlédněme si $a^{\lambda/2}$ přes izomorfismus

$$b\colon \mathbb{Z}_N^* \to \mathbb{Z}_{p_1^{e_1}}^* \times \mathbb{Z}_{p_2^{e_2}}^* \times \cdots \times \mathbb{Z}_{p_k^{e_k}}^*$$

Musí platit $b(a^{\lambda/2}) = (\pm 1, \pm 1, \ldots, \pm 1)$, kde $-1 \in \mathbb{Z}_{p_i^{e_i}}^*$ značí prvek $p_i^{e_i} - 1$. Protože $a \notin K$, musí být nějaká ze souřadnic rovna jedné a nějaká jiná rovna mínus jedné.
Pak $\mathrm{NSD}(a^{\lambda/2} - 1, N)$ bude vlastní dělitel N.
Problém k dořešení: Neznáme λ.

Klíčové lemma

Pro každé $a \in \mathbb{Z}_N^* \setminus K$ a pro každé $x \in \mathbb{N}$ splňující $a^x \equiv 1 \pmod N$ existuje $k \in \mathbb{N}_0$, $0 \leq k \leq \log(x)$ tak, aby

$$a^{x/2^k} \not\equiv \pm 1 \pmod N \quad \& \quad (a^{x/2^k})^2 \equiv 1 \pmod N .$$

Důkaz: Označme $\alpha := o(a)$. Protože je $a^\lambda \equiv 1 \pmod N$, bude $\alpha \mid \lambda$. Dále musí platit $\alpha \nmid \frac{\lambda}{2}$, jinak by bylo $a^{\lambda/2} \equiv 1 \pmod N$, to pro $a \notin K$ není možné. Musí proto platit $\lambda = \alpha\beta_0$, kde $\beta_0 \in \mathbb{N}$ je liché.
Z předpokladu $a^x \equiv 1 \pmod N$ máme $\alpha \mid x$. Nechť $x = \alpha 2^\nu \beta_1$, kde $\nu \in \mathbb{N}_0$ a $\beta_1 \in \mathbb{N}$ je liché.
Pokud $c \in \mathbb{Z}_N^*$ splňuje $b(c) = (\pm 1, \pm 1, \cdots, \pm 1)$ (taky se mohlo napsat $c^2 \equiv 1 \pmod N$), platí $b(c) = b(c^\ell \bmod N)$ pro každé liché $\ell \in \mathbb{N}$.
Proto $b(a^{\lambda/2} \bmod N) = b(a^{\alpha/2} \bmod N) = b(a^{x/2^{\nu+1}} \bmod N)$.
Položíme $k := \nu + 1$. Pak $a^{x/2^k} \not\equiv \pm 1 \pmod N$ přitom $a^{x/2^{k-1}} \equiv 1 \pmod N$. Dále je $1 \leq \nu + 1 \leq \log(x)$ (α je sudé.)

Nalezení exponentu

Využijeme jednoduché pozorování: Pokud $p \in \mathbb{P}$ splňuje $p \nmid o(a)$, jsou podgrupy $\mathbb{Z}_N^*$ generované prvky a a a^p stejné. Pokud umíme vyřešit zobecněný problém diskrétního logaritmu v cyklických podgrupách $\mathbb{Z}_N^*$, najdeme $y \in \mathbb{N}$ tak, aby $(a^p)^y \equiv a \pmod{N}$. Pak stačí položit $x := py - 1$ a dostaneme $a^x \equiv 1 \pmod{N}$.
Snadno si rozmyslíme, že mezi prvními $\lceil \log(\varphi(N)) \rceil$ prvočísly najdeme nějaké p, pro které platí $p \nmid \varphi(N)$. Pak také $p \nmid o(a)$ pro každé $a \in \mathbb{Z}_N^*$.

Popis faktorizačního algoritmu

Zvolíme $a \in \mathbb{Z}_N^*$, doufáme, že $a \notin K$.
Postupně budeme procházet prvočísla $p \in \{2, 3, 5, 7, \dots\}$ a zkoušet, zda orákulum vrátí $y \in \mathbb{N}$ splňující $(a^p)^y = a$.
Nemáme specifikováno, co bude orákulum dělat v případě nekorektního vstupu, tj, pokud $a \notin \langle a^p \rangle$. Můžeme například předpokládat, že v takovém případě sdělí, že úloha nemá řešení.
Nakonec projdeme mocniny dvou dělící $x := py - 1$ a zkusíme výpočtem $\mathrm{NSD}(a^{x/2^k} - 1, N)$ najít vlastní dělitel N.

Algoritmus

Vstup: $N \in \mathbb{N}$ liché, dělitelné alespoň dvěmi různými prvočísly

Výstup: Vlastní dělitel N, nebo neúspěch

k dispozici: Orákulum pro zobecněný DLP pro podgrupy $\mathbb{Z}_N^*$.

1. Vol $a \in \mathbb{Z}_N^*$ náhodně

2. **for** $p \in \{2, 3, 5, 7, \ldots, p_{\lceil \log(N) \rceil}\}$ **do**

Zkus najít $y \in \mathbb{N}$ splňující $(a^p)^y \equiv a \pmod{N}$

Pokud y nalezeno, $x := py - 1$, **break**

3. $k := 0$

while $(2^{k+1} \mid x)$ a $(a^{\frac{x}{2^{k+1}}} \equiv 1 \pmod{N})$ **do** $k := k + 1$

4. $d := 1$

if $2^{k+1} \mid x$ **then** $d := \mathrm{NSD}(a^{\frac{x}{2^{k+1}}} - 1 \bmod N, N)$.

5.if $1 < d < N$ **then return** d

else return neúspěch

Pár slov ke složitosti

Umocnění typu $a^x \bmod N$ provedeme v čase $O(\log(x)\log^2(N))$.
Abychom mohli algoritmus považovat za polynomiální potřebujeme rozumně velké x.
Můžeme například požadovat, aby výstup orákula byl vždy omezen $y \leq N^e$, kde e je nějaká konstanta.
Počet prvočísel pro iterace kroku 2. je pak $\leq \log(N) + 1$, dále máme

$$\lceil \log(N) \rceil \geq cp_{\lceil \log(\mathrm{N}) \rceil} / \ln(p_{\lceil \log(N) \rceil})$$

dává odhad na hodnotu $p_{\lceil \log(N) \rceil}$
Počet iterací v kroku tři je omezený $\log(x) + 1$.

Index calculus pro $\mathbb{Z}_p^*$

Předpokládejme, že $G = \mathbb{Z}_p^*$, kde $p \in \mathbb{P}$ známe generátor g a pro $h \in G$ chceme spočíst $\mathrm{dlog}_g(h)$.
Ukážeme, si princip negenerického algoritmu, který využívá následující vlastnosti G.

- Prvky G jsou přirozená čísla.
- Násobení v G je pro některé prvky totožné s násobením v $\mathbb{Z}$. Přesněji pro $a, b \in G$ takové, že $a._{\mathbb{Z}} b < p$, platí $a._{\mathbb{Z}} b = a._G b$

První fáze algoritmu

Tuto fázi lze chápat jako jakýsi předvýpočet, který stačí provést pro každé prvočíslo pouze jednou.
Zvolíme faktorizační bázi $B = \{2, 3, 5, \ldots, p_{max}\}$. Cílem první fáze je určení $\mathrm{dlog}_g(b)$ pro každé $b \in B$.
Postup je docela jednoduchý: Zvolíme náhodně $r \in \mathbb{Z}_{p-1}$ a spočteme $g_r := g^r \pmod p \in \mathbb{Z}_p^*$.
Prvek g_r je celé číslo, pokud není součinem prvků B (tedy pokud g_r není p_{max}-hladké), volíme nové r.
V opačném případě najdeme faktorizaci $g_r = \prod_{b \in B} b^{e_{r,b}}$, kde $e_{r,b} \in \mathbb{N}_0$.
Tuto rovnost lze také chápat jako rovnost v $\mathbb{Z}_p^*$. Jejím zlogaritmováním dostaneme

$$\sum_{b \in B} e_{r,b} \mathrm{dlog}_g(b) \equiv r \bmod (p-1)$$

První fáze, dokončení

Postup opakujeme tak dlouho, dokud nenasbíráme tolik kongruencí, že jejich soustava bude mít v $\mathbb{Z}_{p-1}^{|B|}$ jediné řešení.
Řešením této soustavy kongruencí dostaneme $\mathrm{dlog}_g(b)$ pro každé $b \in B$.
Je potřeba si uvědomit, že se nejedná o soustavu lineárních rovnic.

Druhá fáze algoritmu

Teprve nyní budeme počítat s prvkem h jehož dlog_g chceme určit. Zvolíme náhodně $r \in \mathbb{Z}_{p-1}$ a spočteme

$$h_r = g^r.h \bmod p\,.$$

Proces opakujeme, dokud h_r není součin prvků B.
Nakonec máme $h_r = \prod_{b\in B} b^{e_b}$. Zlogaritmováním obdržíme

$$\mathrm{dlog}_g(h_r) \equiv r + \mathrm{dlog}_g(h)(\mathrm{mod}(p-1))$$

$$\mathrm{dlog}_g(h_r) \equiv \sum_{b\in B} e_b \mathrm{dlog}_g(b) \ (\mathrm{mod}\ (p-1))\,.$$

Hodnoty $\mathrm{dlog}_g(b)$ známe z první fáze algoritmu. Proto

$$\mathrm{dlog}_g(h) = (-r + \sum_{b\in B} e_b \mathrm{dlog}_g(b)) \bmod (p-1)$$

Poznámka ke složitosti

Průměrnou dobu běhu lze odhadnout podobnými metodami, kterými jsme odhadovali složitost kvadratického síta. Tentokrát nás zajímá s jakou pravděpodobností je náhodně volený prvek $\mathbb{Z}_p^*$ $p_{\max}$ hladký. Asymptotické chování lze určit z de Bruijnovy funkce. Při optimalizaci volby B, testování hladkosti lze očekávat, že průměrná složitost první fáze je $\mathrm{L}_p[\frac{1}{2}, \sqrt{2} + o(1)]$ a průměrná složitost druhé fáze je $\mathrm{L}_p[\frac{1}{2}, \frac{1}{\sqrt{2}} + o(1)]$.

Konec

Děkuji za pozornost.